

Отзыв руководителя на выпускную квалификационную работу

обучающегося СПбГУ Бзикадзе Андрея Важевича

Статистические свойства некоторых процедур сжатия данных

Выпускная магистерская работа Андрей Бзикадзе является продолжением его бакалаврской работы, защищенной два года назад и посвященной теоретическому анализу статистического теста «Book Stack», предложенного специалистами по сжатию данных для проверки «нулевой» статистической гипотезы о том, что наблюдаемые данные являются повторной выборкой из дискретного равномерного распределения с известным носителем. Общая идея метода, предложенного его авторами, может быть описана так: исходная выборка подвергается некоторому преобразованию «Book Stack», после чего именно преобразованная выборка (а не исходная) подвергается статистическому анализу.

В бакалаврской работе А. Бзикадзе (см. также опубликованную по ее следам статью в Вестнике СПбГУ) показано, что подобная процедура, вообще говоря, приводит к уменьшению мощности в случае, если альтернативной гипотезе соответствует повторная, но не равномерно распределенная выборка.

В настоящей работе рассматривается другая (и более адекватная по отношению к проверке качества генераторов псевдослучайных чисел) альтернатива. А именно, предполагается, что исходная выборка является реализацией эргодической однородной марковской цепи, стационарное распределение которой является равномерным. Следовательно, здесь проверяется не отсутствие (асимптотической) равномерности распределенности, а отклонение от независимости.

В этой ситуации многие стандартные статистические процедуры проверки нулевой гипотезы не являются, естественно, состоятельными.

В магистерской работе показано на примере стандартного критерия хи-квадрат, что после преобразования данных критерий становится состоятельным, если только след переходной матрицы не равен нулю. Как показывают приведенные примеры, при нарушении этого условия состоятельность может нарушиться.

Этот результат, использующий законы больших чисел, можно считать центральным, если привязываться только к вышеописанной статистической постановке задачи.

Но, кроме того, в работе получены и другие интересные результаты.

Так, например, там доказана центральная предельная теорема для выходной последовательности «Book Stack», рассматриваются обобщения этого преобразования; есть и другие, в основном промежуточные результаты, а также анализ еще одного статистического «Order» теста, основанного на аналогичной процедуре.

Имеет смысл отметить большую самостоятельность Андрея. Собственно, роль руководителя здесь скорее сводилась к обсуждению полученных результатов, чем к постановке задач и проверке доказательств.

Работа удовлетворяет всем критериям, предъявляемым к выпускной квалификационной работе в магистратуре, и заслуживает оценки «отлично».

Научный руководитель, доцент кафедры статистического моделирования

23.05.2018



В.В. Некруткин